Brain Research 1646 (2016) 475–481

Contents lists available at ScienceDirect

Brain Research

journal homepage: www.elsevier.com/locate/brainres

Research report

Incremental comprehension of spoken quantifier sentences: Evidence from brain potentials

Dominik Freunberger [a,*], Mante S. Nieuwland [b]

[a] Centre for Cognitive Neuroscience, University of Salzburg, Salzburg, Austria
[b] Department of Psychology, School of Philosophy, Psychology and Language Sciences, University of Edinburgh, Edinburgh, United Kingdom

ARTICLE INFO

Article history:
Received 23 December 2015
Received in revised form
25 May 2016
Accepted 23 June 2016
Available online 23 June 2016

Keywords:
Quantifier
Incremental interpretation
Prediction
Spoken language comprehension
ERP
N400

ABSTRACT

Do people incrementally incorporate the meaning of quantifier expressions to understand an unfolding sentence? Most previous studies concluded that quantifiers do not immediately influence how a sentence is understood based on the observation that online N400-effects differed from offline plausibility judgments. Those studies, however, used serial visual presentation (SVP), which involves unnatural reading. In the current ERP-experiment, we presented spoken positive and negative quantifier sentences ("Practically all/practically no postmen prefer delivering mail, when the weather is good/bad during the day"). Different from results obtained in a previously reported SVP-study (Nieuwland, 2016) sentence truth-value N400 effects occurred in positive and negative quantifier sentences alike, reflecting fully incremental quantifier comprehension. This suggests that the prosodic information available during spoken language comprehension supports the generation of online predictions for upcoming words and that, at least for quantifier sentences, comprehension of spoken language may proceed more incrementally than comprehension during SVP reading.

1. Introduction

Quantifier expressions like 'most' or 'few' are crucial for people to communicate information about the world in an efficient manner. People use quantifiers to express to what extent a certain property holds true for the entities belonging to a larger set (e.g. "Most whiskies in Edinburgh pubs are Scotch"). In philosophical and linguistic theories of meaning, quantifiers therefore are commonly associated with their impact on sentence truth-conditions (e.g. Partee, 1991), the conditions that make a sentence ultimately true or false. Whereas truth-conditions are typically considered without regard to how the sentence unfolds in time when people listen to or read a sentence, quantifier meaning may influence the listener's comprehension of the unfolding sentence before it is finished. A question thus arises: Do people incrementally incorporate the meaning of quantifier expressions to understand an unfolding sentence? Previous research has sought an answer to this question by examining whether and when on-line measures of comprehension (e.g., eye-movements or ERPs) correspond to the meaning that readers ultimately extract from quantifier sentences as reflected in offline measures (e.g., plausibility or truth-value judgments). This research typically examines whether truth-value or plausibility impacts online comprehension in positive and negative quantifier sentences alike. Most studies to date report that on-line and offline measures show different patterns of quantifier comprehension, suggesting that quantifier meaning does not fully incrementally impact the interpretation of an unfolding sentence (e.g., Kounios and Holcomb, 1992; Urbach and Kutas, 2010). However, previous studies examined comprehension of written language, using artificial reading procedures such as serial visual presentation. It is an open question whether people incorporate quantifier meaning in an incremental manner when listening to spoken sentences. The present study investigated this issue by recording neurophysiological responses to spoken quantifier sentences.

1.1. Quantifiers and incremental interpretation

Previous ERP research on quantifier comprehension has investigated the processing consequences of sentences that do not correspond to what people hold to be true or plausible in the real world (Kounios and Holcomb, 1992; Nieuwland, 2016; Urbach et al., 2015; Urbach and Kutas, 2010). The dependent measure in these studies is the N400 ERP component (Kutas and Hillyard, 1980), a negative ERP deflection peaking around 400 ms after word-onset. N400 amplitude is smaller when the retrieval of word-associated information in semantic memory is facilitated by

* Correspondence to: Centre for Cognitive Neuroscience, Department of Linguistics, University of Salzburg, Erzabt Klotz Strasse 1, 5020 Salzburg, Austria.
E-mail address: dominik.freunberger@sbg.ac.at (D. Freunberger).

http://dx.doi.org/10.1016/j.brainres.2016.06.035
0006-8993/

the context (e.g., Kutas and Federmeier, 2011), potentially via pre-activation of relevant information (e.g., Ito et al., 2016). Words that render a sentence true elicit a smaller N400 than words that render a sentence false (Nieuwland and Kuperberg, 2008; Nieuwland and Martin, 2012), reflecting the facilitated comprehension of words that render a sentence true. In research on quantifier comprehension, the question of interest is whether such sentence truth-value N400 effects occur in positive and negative quantifier sentences alike.

In the first ERP study on quantifier comprehension, Kounios and Holcomb (1992) found no effect of quantifier type (positive or negative) on the N400 to the last words of sentences like "All/No rubies are gems/spruces", even though the sentences were evaluated correctly after they were finished. Kounios and Holcomb concluded that quantifier interpretation is delayed and that initial semantic processes as indexed by the N400 are insensitive to the compositional meaning of the sentence, and only reflect lower-level lexical-associative relationships (cf. Fischler et al., 1983). However, the results are also consistent with a step-wise account of sentence verification, in which readers initially compute and evaluate an affirmative proposition before applying negative quantifier meaning (Carpenter and Just, 1975).

A different pattern was observed by Urbach and Kutas (2010). In positive and negative quantifier sentences (e.g., "Almost all/Almost no groupies follow singers/boys"), atypical objects like 'boys' elicited the same N400 regardless of quantifier type, whereas typical objects like 'singers' elicited smaller N400s following positive quantifiers compared to negative quantifiers. The authors took these results as evidence for partial incremental comprehension of negative quantifiers, as the online N400 measures did not mirror the post-sentence plausibility ratings (atypical objects were judged less plausible than typical objects in positive sentences, but more plausible in negative sentences).

In a very recent study by Urbach et al. (2015), this pattern of results changed to a more incremental pattern (i.e., smaller N400 for implausible sentences regardless of quantifier type) when a supportive discourse preceded the quantifier sentences (e.g., "Alex was an unusual toddler. Few/Most children prefer vegetables/sweets"). However, this incremental pattern occurred *only* when participants were not required to explicitly evaluate plausibility of the sentences. A partial incremental pattern similar to that of Urbach and Kutas (2010) was observed when participants made plausibility judgments following each sentence. The authors concluded that task variables appear to impact the speed and/or depth of quantifier interpretation, although their discussion fell short of an explanation for why quantifier interpretation would be slower or less deep when people are engaged in a task that explicitly probes quantifier sentence meaning.

To account for the different patterns observed in previous studies, Nieuwland (2016) recently proposed a prediction-based account of online quantifier interpretation: the precise pattern in which quantifiers are understood depends on the extent to which quantifier meaning is incorporated into a prediction for upcoming words. This proposal was based on an ERP study wherein participants read sentences such as "Practically all/no postmen prefer delivering mail when the weather is good/bad". When positive and negative quantifiers had similarly low cloze-values, a pattern comparable to that reported by Kounios and Holcomb (1992) was observed, whereas when positive and negative quantifiers had similarly high cloze-values, sentence truth-value N400-effects occurred regardless of quantifier type (i.e., a fully incremental pattern; see also Nieuwland and Martin (2012)). Quantifier sentences are thus understood neither always in two sequential stages, nor always in a partial-incremental fashion, nor always in a maximally incremental fashion. Fully incremental quantifier interpretation only occurs when quantifiers are incorporated into sufficiently strong online predictions for upcoming words.

The prediction-based account proposed by Nieuwland (2016) captures the range of previous results on online quantifier comprehension, which all involved reading. In reading ERP studies, words are typically presented one at a time in the middle of the screen (serial visual presentation, SVP), at a fixed pace. While SVP has the benefit of minimizing eye-movement artefacts during EEG recording, this procedure is unnatural because it does not allow participants to read at their own pace and to preview upcoming information. Importantly, the ability to generate online predictions may be limited by SVP. One relevant factor is the word presentation duration. Uncomfortably fast-paced SVP may make it difficult for people to generate predictions even when they read relatively high-cloze sentences (e.g., Dambacher et al., 2012), whereas slow-paced SVP may allow people more time to generate predictions even in low-cloze sentences. Interestingly, Kounios and Holcomb (1992), who found no effect of quantifier type on the N400, had a more rapid presentation rate per word on average (850 for 3 words such as "No dogs are"). A second important factor is that spoken sentences contain rich prosodic information that aids comprehension (Cutler et al., 1997; Frazier et al., 2006), which may facilitate the online generation of predictions. In the present study, we therefore examine the comprehension of spoken quantifier sentences.

1.2. Incrementality in spoken language comprehension

Spoken language comprehension typically proceeds in a highly incremental manner: at each moment in time, listeners effectively use the smallest cues that are available (phonemes) to interpret the unfolding sentence and sometimes to generate predictions about what comes next (Altmann and Mirković, 2009). The interpretation of a spoken word therefore starts as early as its first phoneme, where lexical candidates that match the input may become activated in parallel until further information narrows down the set of candidates to one uniquely identifiable word (e.g. Marslen-Wilson, 1987). Moreover, ERP studies have shown that when listeners generate a prediction about which word comes next, the first bit of input that is inconsistent with the predicted input elicits processing costs as reflected in the N400 (Van Petten et al., 1999).

These online predictions about upcoming words are based on the meaning of the unfolding sentences, but they are also strengthened by co-articulation, i.e., the assimilation of the pronunciation of two neighbouring words (Öhman, 1966). The pronunciation of one word thus typically provides prosodic information about the next word. Listeners can benefit from co-articulation in similar ways as readers benefit from parafoveal preview of upcoming words during natural reading: both phenomena will facilitate recognition and comprehension of upcoming words (Rayner, 1998). Co-articulation might therefore contribute to the relatively earlier N400 onset in spoken language compared to SVP (Kutas et al., 1987; Van Berkum et al., 2003; but, see also Hagoort and Brown (2000a)). This earlier onset is only observed in spoken sentences, because when the words of a spoken sentence are presented at a fixed rate, the N400 time-course is more comparable to SVP reading (Holcomb and Neville, 1991). In addition to co-articulation, comprehension is also facilitated by sentence prosody, as listeners are able to use prosody to predict utterance length and determine phrase boundaries (e.g. Cutler et al., 1997).

The comprehension of spoken words is thus strongly influenced both by the prosodic and linguistic context in which they appear. Compared to SVP, where a word presented as a whole confirms or disconfirms a prediction, listeners can determine whether or not the word matches the prediction based on only a tiny bit of spoken input. Also compared to SVP reading, listeners benefit from co-articulation, as more information is available to generate a prediction of the next word. Because predictive

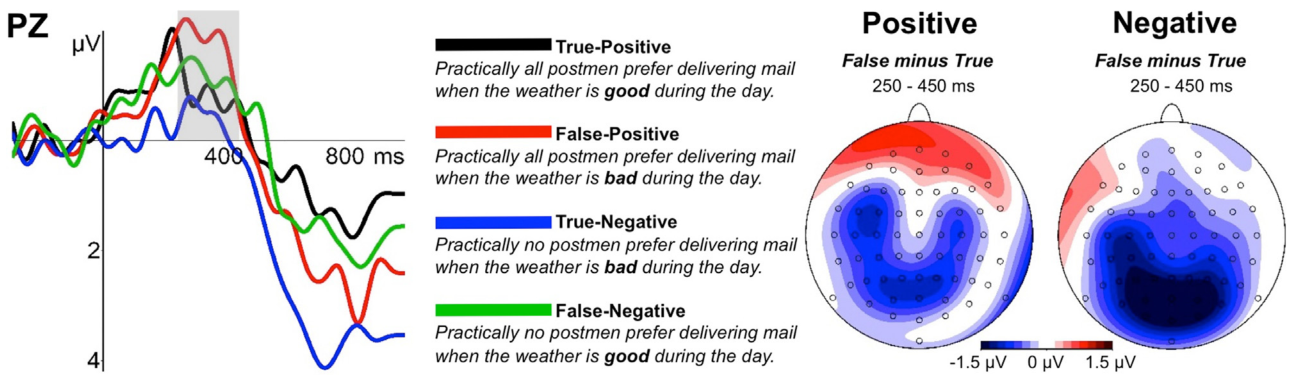

Fig. 1. Grand average ERPs at PZ, time-locked to the onset of the CW in true and false sentences containing positive and negative quantifiers. The grey area corresponds to the time-window used for N400 analyses. For illustration only, data was low-pass filtered at 10 Hz. An example sentence for each condition is given in the middle (CWs are bold). On the right side, difference maps (false minus true) for positive and negative quantifier sentences illustrate the distribution of the N400 truth-effect.

processing may be stronger during spoken language comprehension than during SVP reading, and because online prediction of upcoming information benefits incremental comprehension of quantifier sentences (Nieuwland, 2016), comprehension of quantifier sentences may proceed more incrementally in listeners than in SVP readers.

1.3. *The present study*

The current experiment aims to examine the time-course of quantifier interpretation in spoken language. We created auditory recordings of sentences from a previous SVP experiment (Nieuwland, 2016), and followed the same experimental logic and procedure. We also performed similar linear mixed-effects model analyses as in the SVP-version of the experiment, with cloze-value and truth-value as continuous predictors, quantifier-type as a fixed factor, and the N400 as dependent variable. If the time-course of quantifier interpretation is comparable in spoken sentences and SVP, our N400 results would replicate the three-way interaction between truth-value, quantifier-type, and cloze-value observed by Nieuwland (2016). In that SVP study, positive and negative quantifier sentences with relatively high cloze values yielded similar truth-value N400-effects, whereas positive and negative quantifier sentences with low cloze values yield different N400-effects, as true words elicit smaller N400s than false words in positive but not negative sentences. Alternatively, the time-course of quantifier interpretation may be more incremental in spoken language comprehension than in SVP reading. N400 truth-value effects may therefore be observed in positive and negative sentences alike, both in low and high cloze sentences.

2. Results

As can be seen in Figs. 1 and 2, true sentences elicited smaller N400s than false sentences: The N400 for true-positive was smaller compared to false-positive ($-.47$ ($SD=9.39$) versus $-.99$ ($SD=9.30$) μV) and true-negative sentences led to a smaller N400 compared to false-negative ones ($.01$ ($SD=10.22$) versus $-.96$ ($SD=9.63$) μV). The early N400 onset is comparable to other auditory studies investigating semantic ERP effects (e.g. Connolly and Phillips, 1994; Hagoort and Brown, 2000b; Holcomb and Neville, 1991; Van Berkum et al., 2003; Van Petten et al., 1999).[1] Moreover, the effect of truth-value appears sustained in negative sentences only. Crucially, CW predictability did not modulate this N400 difference between true and false sentences: the LME-model that included the three-way interaction between cloze-value, truth-value and quantifier-type (model 1) did not improve fit over the model without this three-way interaction (model 2), $\chi^2(1)=1.09$; $p=.3$.[2]

In model 2, the smaller N400 for true compared to false sentences was reflected in a main effect of truth-value ($t(1)=2.14$), independent of quantifier-type ($t<1$). Additionally, there was a main effect of cloze-value ($t(1)=2.83$), reflecting the smaller N400s for high-cloze compared to low-cloze sentences. Cloze-value did not interact with quantifier-type or truth-value (all ts <1). Analysis in the 500–800 ms time-window yielded a reliable effect of truth-value ($t=2.01$) as well as a quantifier-type by truth-value interaction ($t=-2.02$). In negative quantifier sentences, the voltage difference between true and false sentences is greater (true minus false, 1.23 μV) than in positive quantifier sentences (true minus false, -0.49 μV).

3. Discussion

In this study, we examined the impact of positive and negative quantifiers on online comprehension of spoken sentences that were either true or false ("Practically all/no postmen prefer delivering mail when the weather is good/bad during the day"). Words that rendered a sentence true elicited smaller N400s than words that rendered a sentence false, which reflects the facilitation of semantic retrieval for words that are consistent with factual world-knowledge. Crucially, we observed this sentence truth-value effect in positive and negative quantifier sentences, suggesting that positive and negative quantifiers did not differ in their impact on online comprehension. These findings are inconsistent with the findings of most previous ERP studies on quantifier comprehension (e.g., Kounios and Holcomb, 1992; Urbach et al., 2015, experiments 1, 3a and 3b; Urbach and Kutas, 2010), which typically showed delayed effects of truth-value or plausibility in negative compared to positive quantifier sentences. Importantly, our results also differ from the overall pattern reported in Nieuwland (2016), which used the same sentence materials but used serial visual presentation. In Nieuwland (2016), a comparable N400 truth-effect in positive and negative quantifiers was only observed when

[1] We created an additional figure from ERPs with a longer baseline period, which is available as an online Supplementary material. This figure shows that the observed effects in the N400 window are not the result of ERP differences before word-onset (i.e., baseline differences).

[2] Model 1: N400~quantifier*truth*cloze +(1+quantifier+truth+cloze | subject)+(1+quantifier+truth | item).

Model 2: N400~quantifier*truth*cloze − quantifier: truth:cloze +(1+quantifier+truth+cloze | subject)+(1+quantifier+truth | item).

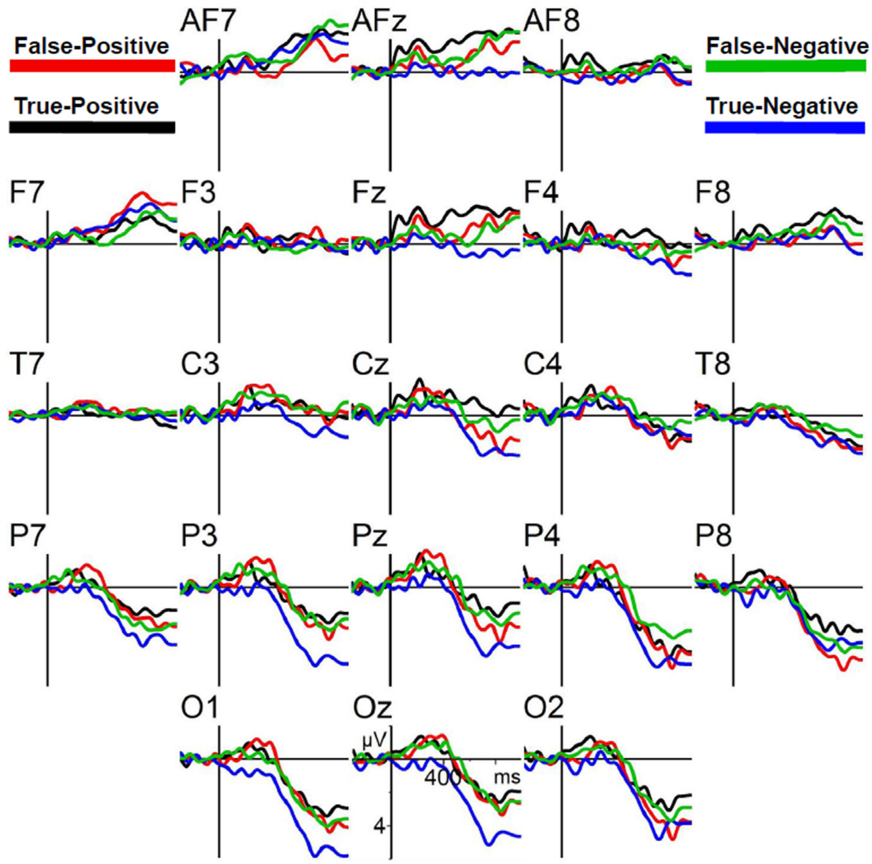

Fig. 2. Grand average ERPs elicited by the critical word from all four conditions at selected electrode sites.

the target words were strongly predictable from their context (i.e., when they had a relatively high cloze-value). Thus, while Nieuwland (2016) reported N400 truth-value effects only in high-cloze sentences, we observed that spoken versions of the same set of materials elicited N400 truth-value effects in high- and low-cloze sentences alike.

To our knowledge, our ERP results are the first to show a different pattern for spoken language comprehension as compared to SVP reading, the standard visual presentation procedure in ERP research on language comprehension. We argue that in the current spoken language study, participants may have been able to use sentence context more effectively in low cloze sentences to anticipate upcoming information. Comprehension of spoken language, at least spoken quantifier sentences, can proceed more incrementally than comprehension of word-by-word SVP reading. We discuss the implications of our findings for accounts of quantifier comprehension and for the use of SVP reading in ERP research.

3.1. *Incremental comprehension of spoken quantifier sentences*

Online quantifier sentence comprehension has been investigated by examining the correspondence between online measurements, such as ERPs, and offline measures that reflect the meaning that people ultimately compute, such as truth-value judgments. Previous accounts of quantifier interpretation differentiated between three possible patterns (see also Urbach et al. (2015)): according to a step-wise account, quantifier meaning does not impact the initial semantic processes reflected in the N400 (Kounios and Holcomb, 1992; cf. Carpenter and Just, 1975). According to a partial-incremental account (Urbach and Kutas, 2010), quantifiers do impact the initial semantic processes reflected in the N400, but do not yield an interpretation that corresponds to the ultimate interpretation as reflected in post-sentence judgments. According to a fully incremental account, which holds that incoming words are related to the widest interpretive background as early as possible (Altmann and Mirkovic, 2009; Carpenter and Just, 1975), the initial semantic processes yield an interpretation corresponding to the final interpretation. Only the latter account thus predicts sentence truth-value N400 effects for negative and positive quantifier sentences alike. Whereas all three accounts have support in the literature, none of them captures the full range of experimental findings. Nieuwland (2016) therefore recently proposed a novel account to accommodate all previous findings. In his prediction-based account, quantifier interpretation is neither always sequential, nor always partial, nor always incremental, but relies on whether people incorporate quantifier meaning into an online prediction for upcoming words. This account was further supported by the observation that with increasing cloze-values the ERP results shifted from a sequential or partial pattern towards a fully incremental pattern.

In contrast to the SVP reading study (Nieuwland, 2016) we observed a full incremental pattern in the entire set of sentences and independent of cloze-value. That is, despite identical content of the materials, participants in our study may have been able to generate relatively stronger predictions, resulting in a facilitation of true words even in negative quantifier sentences with lower cloze-values. The rich prosodic information, which is absent in SVP reading, seems to strengthen listeners' expectations about upcoming words (cf. Brown et al., 2011) and, consequently, leads to increased incremental processing compared to SVP reading.

However, we do not think that all previous quantifier studies would have observed a full incremental pattern with spoken sentences. More naturalistic speech might strengthen predictions relative to SVP reading, but only when sentences allow some amount of semantic prediction in the first place. As argued by Nieuwland (2016), this may not have been the case in previous studies that used low-constraint sentences. Even in spoken sentences, the predictability of unrelated objects such as 'cigarettes' following "Almost no lifeguards protect" remains low, simply because the sentence context does not effectively narrow down the almost infinite number of things that can follow. We therefore predict that for spoken sentences such as "Almost no/all lifeguards protect birds/swimmers" (Urbach et al. 2015), the observed pattern of results may be very similar to the observed patterns in SVP reading. Nevertheless, the current results suggest that auditory presentation of spoken sentences can enhance people's predictions and – under certain conditions – can lead to qualitatively different results than the commonly employed SVP.

A caveat to our conclusion regarding the similarity with which participants processed positive and negative quantifier sentences is that differences occurred in the later, post-N400 time-window. In this time-window, the voltage difference between true and false negative quantifier sentences was larger than the difference between true and false positive quantifier sentences. This may reflect an extended N400-effect in the negative quantifier sentences beyond the traditional 300–500 ms statistical testing-window, which has been observed in previous ERP-studies (e.g., Nieuwland and Van Berkum, 2006; Van Berkum et al., 2003). However, an alternative interpretation is that it reflects the effects of a strongly positive-going later ERP waveform in the true-negative quantifier sentences. The latter interpretation is more consistent with the fact that the differential effect is numerically greater in the 500–800 ms time-window than the N400 time-window. Such a positive ERP-effect may well be associated with the task demands. In the current study, all participants performed a post-sentence verification task, and such tasks are known to elicit positive going waveforms that are probably related to task-related decision making (see also Nieuwland (2015, 2016), who shows ERP waveforms for the same materials when participants are doing verification and when they are not; in both studies the verification instruction was associated with positive-going ERP waveforms). We do not have a good explanation for why the true-negative condition specifically elicited such a strong positive ERP waveform, one possibility is that they paid extra attention to negative quantifier sentences. Future studies can address this issue by testing the impact of truth-value in spoken quantifier sentences when participants are not engaged in a secondary verification task.

3.2. Implications for SVP research on language comprehension

In EEG/MEG and fMRI research on language comprehension, researchers typically use serial visual presentation (SVP). SVP is easier to implement than spoken language, and has the benefit over natural reading that observed brain responses can be straightforwardly linked to word onset. Most SVP research assumes, either explicitly or implicitly, that language comprehension processes proceed similarly in SVP compared to listening or natural reading. Our current data, however, call this assumption about the representativeness of SVP reading into question.

We argue that when researchers ask a question about the time course of comprehension, such as in research on incremental and predictive processing, SVP reading may distort the picture of what goes on during more naturalistic language comprehension. We think that this may be particularly true for syntactically or semantically complex sentences (e.g., object-relative sentences, negation sentences), while perhaps less true for sentences with relatively simple affirmative sentences especially when they involve a semantic anomaly. Previous comparisons of SVP and spoken language comprehension have compared ERP responses to strong semantic or syntactic anomalies (e.g., Hagoort and Brown, 2000a; Kutas et al., 1987), which inherently involve unexpected sentence anomalies and where online predictions do not generate the main effect of interest. As discussed in the previous section, we think that the differences in incremental processing between SVP reading and spoken language comprehension (and possibly natural reading) are most likely to surface when the ability to generate online predictions does matter. When online predictions can in principle be formed, the chances of this happening may be greater during spoken language comprehension than during SVP reading. As a consequence, absence of prediction effects or evidence for a delay in interpretation may partially be the result of SVP reading. General conclusions about incremental processing solely based on SVP results can therefore be misleading.

SVP reading does not allow the pre-view that can occur during natural reading, and lacks the prosody and co-articulation of spoken language. However, the general speed of the SVP presentation rate may also matter. Most studies use a standard presentation rate of approximately two words per second (for an alternative approach, see, e.g., Legge et al., 1997). This conventional rate was not initially determined for a specific reason other than being a rate at which people read comfortably. While its relatively slow pace makes up for the fact that readers cannot benefit from preview, perhaps even this relatively slow pace can weaken expectations for upcoming words. Consistent with this interpretation, several SVP studies have reported that online prediction effects are stronger at slower rates (e.g., Ito et al., 2016; cf. Dambacher et al., 2012). A systematic comparison of different input rates, of written and spoken language, is warranted to determine their effects on predictive processing. We suspect that input rate has a greater effect in SVP reading than in spoken language comprehension and that perhaps at a slower SVP presentation rate the patterns of quantifier comprehension are more similar to how spoken quantifier sentences are understood.

3.3. Conclusion

Incremental interpretation of linguistic input is regarded as one of the major characteristics of human language comprehension (Altmann and Mirković, 2009). Results from previous quantifier studies have challenged the notion that language comprehension is fully incremental and relates incoming input to the widest interpretive background at the earliest possibly moment. Using the serial visual presentation procedure, sentence truth-value N400 effects are often observed for positive quantifier sentences but not for negative quantifier sentences (e.g. Kounios and Holcomb, 1992). The current study on comprehension of spoken quantifier sentences yielded two novel insights. First, the comprehension of spoken sentences proceeds fully incrementally, as suggested by N400 truth-value effects in positive and negative quantifier sentences alike. This finding adds to the recent reports that negative quantifier sentences are not inherently more difficult to understand than positive quantifier sentences (Nieuwland, 2016; Urbach et al., 2015, experiment 2). Second, people use spoken language more effectively to generate online predictions about upcoming words than written SVP input. Consequently, comprehension may typically proceed more incrementally during listening to natural speech than during SVP reading.

4. Methods

4.1. Participants

Participants were thirty right-handed English native speakers (21 female), aged between 19 and 39 years (mean 24 years), without history of neurological or psychiatric disorders, who had not participated in Nieuwland (2016). Written informed consent was obtained and participants were paid for the duration of the experiment.

4.2. Materials

Materials consisted of 124 sentence-quadruplets, which started either with a positive or a negative quantifier expression, and which contained a CW that rendered each sentence true or false dependent on the quantifier (e.g. "Practically no/all postmen prefer delivering mail when the weather is good/bad during the day"). CWs that rendered positive quantifier sentences true rendered the negative quantifier sentence false, and vice versa. For description of how the sentences were created and for a full list of the materials, see Nieuwland (2016).

The sentences were recorded with a normal speaking rate of approximately 2.1 words per second (cf. Tauroza and Allison, 1990) and normal intonation by a female British native speaker at a sampling rate of 48 kHz. Only the true-positive and true-negative version of each item was recorded, to avoid prosodic or other differences between false and true sentences. False sentences were created by splicing the beginning of each sentence so that the cut was not audible, which was between 1 and 5 words after the quantifier expression, and substituting the negative quantifier phrase for the positive quantifier phrase and vice versa. Average critical word duration was 624 ms (ranging from 179 to 1071 ms) and the average time between the onset of the quantified noun and the critical word was 3759 ms (ranging from 2313 to 5302 ms). Of the 124 critical word pairs, only 9 pairs shared the initial phoneme.

4.3. Procedure

Participants were seated in a sound-attenuating booth in front of a computer monitor and sentences were presented over speakers placed left and right of the monitor. Participants were instructed to closely attend to the sentences and to avoid eye and other movements while listening. After each sentence, a response display with the response options "1-2-3-4-5" (with "Strongly disagree" under the 1 and "Strongly agree" under the 5) was shown und participants were asked to respond according to whether or not they agree with the sentence they just heard by pressing the according button on a keyboard. There was no time limit and participants were instructed to respond as accurately as possible. Only trials with condition-consistent responses (4 and 5 for true, 1 and 2 for false sentences) were included in the analyses, which left an average of 86% of all trials. The experiment was divided into seven blocks, between which participants could make short breaks. Total experiment time was approximately 70 min.

4.4. EEG recording and data processing

EEG was recorded at 512 Hz sampling rate – from 64 scalp electrodes and an additional four EOG and two mastoid electrodes using a BioSemi Active Two system (for electrode locations, see the voltage maps in Fig. 1). Data was recorded against an active electrode reference (common mode sense) and a passive electrode ground. Scalp impedances were kept below 5 kΩ. Note that the recording was identical to Nieuwland (2016). Offline, data was band-pass filtered from .019 to 20 Hz and segmented into epochs from −500 to +1000 ms around critical word onset. Ocular artefacts were corrected using the Gratton and Coles method before data was baseline-corrected to −200 to 0. Based on an exclusion criterion of more than 40% of trial loss due to artefacts and condition-inconsistent responses, 3 participants were excluded from further analysis. From the remaining 27 participants, an average of 78% of the trials entered analyses[3] (12% loss due to condition inconsistent responses and 10% due to EEG artefacts).

4.5. Statistical analysis

Following Nieuwland (2016), we performed mixed-effect model analysis (LME; Baayen et al., 2008) using the lme4-package (Bates et al., 2012) in R (R Core Team, 2014). The dependent variable was N400-amplitude per item in a 250–450 ms time-window averaged across 22 posterior electrodes (CP1/3/5, P1/3/5/7, TP7, O1, PO3/7 and right equivalents; similar to Nieuwland, 2015)[4] This time window best captured the observed N400 component, which is often earlier in auditory experiments than in visual experiments (e.g., Kutas and Federmeier, 2011). Additionally, we performed comparable LME-analyses in a 500–800 ms time window in order to capture possible effects that go beyond the N400 time-window.

We defined a model that included the three-way interaction between quantifier type (positive, negative), truth-value (the response to each item in the post-sentence verification task) and cloze probability (the mean cloze value for true-positive and true-negative sentences) as a fixed effect. Random effects (intercepts for subjects and items) were quantifier-type, truth-value, and cloze in the by-subject slope (without their interaction), and quantifier-type and truth-value in the by-item slope (again, without their interaction; see footnote 2 for model definitions). We opted for this model, because models that included the full interaction terms in the random slopes did not converge (cf. Barr et al., 2013). The resulting model (model 1) was compared to a second model with the same predictors but without the three-way interaction between cloze-value, truth-value, and quantifier type (model 2) by employing a likelihood ratio test using an ANOVA. Resulting *t*-values of 2.00 and above were treated as significant.

Acknowledgements

The first author is financially supported by the Doctoral College "Imaging the Mind" of the Austrian Science Fund (FWF- W1233). This work was funded by British Academy grant SG131266 to MSN.

Appendix A. Supplementary material

Supplementary data associated with this article can be found in the online version at http://dx.doi.org/10.1016/j.brainres.2016.06.035.

[3] Our artifact rejection procedure led to a different number of participants in the four counterbalanced stimulus lists. However, we should note that our mixed-effects model analysis approach, by modelling item-level variance in addition to subject-level variance, deals straightforwardly with list imbalances and also generates more robust effects when there are fewer observations than analysis based on subject-averages as often done in ERP research (cf. Baayen et al., 2008).

[4] To verify the posterior distribution of the N400, we pooled 44 electrodes into four quadrants, each with 11 channels: Anterior-left/right (FP1/2, AF3/4, AF7/8, F1/2, F3/4, F5/6, F7/8, FC1/2, FC3/4, FC5/6, FC7/8), Posterior-left/right (CP1/2, CP3/4, CP5/6, TP7/8, P1/2, P3/4, P5/6, P7/8, PO3/4, PO7/8, O1/2). Using the distributional factors anteriority (anterior, posterior) and hemisphere (left, right) allowed testing for anterior-posterior as well as hemispheric differences. The mean voltages from 250 to 450 ms per condition and participant were submitted to a 2(quantifier-type: negative, positive) by 2(truth-value: true, false) by 2(anteriority: anterior, posterior) by 2(hemisphere: left, right) analysis of variance. The ANOVA yielded a significant truth-value by anteriority interaction ($F(1,26)=5.22$, $p<.031$). Resolution of this interaction revealed that there was a reliable effect of truth-value at posterior ($F(1,26)=9.77$, $p<.005$) but not at anterior regions ($F<1$).

References

Altmann, G., Mirković, J., 2009. Incrementality and prediction in human sentence processing. Cognit. Sci. 33, 583–609.
Baayen, R.H., Davidson, D.J., Bates, D.M., 2008. Mixed-effects modeling with crossed random effects for subjects and items. J. Mem. Lang. 59, 390–412.
Barr, D.J., Levy, R., Scheepers, C., Tily, H.J., 2013. Random effects structure for confirmatory hypothesis testing: keep it maximal. J. Mem. Lang. 68, 255–278.
Bates, D., Maechler, M., Bolker, B., 2012. lme4: Linear Mixed-Effects Models Using S4 Classes.
Brown, M., Salverda, A.P., Dilley, L.C., Tanenhaus, M.K., 2011. Expectations from preceding prosody influence segmentation in online sentence processing. Psychon. B. Rev. 18, 1189–1196.
Carpenter, P.A., Just, M.A., 1975. Sentence comprehension: a psycholinguistic processing model of verification. Psychol. Rev. 82, 45–73.
Connolly, J.F., Phillips, N., 1994. Event-related potential components reflect phonological and semantic processing of the terminal word of spoken sentences. J. Cognit. Neurosci. 6, 256–266.
Cutler, A., Dahan, D., Van Donselaar, W., 1997. Prosody in the comprehension of spoken language: a literature review. Lang. Speech 40, 141–201.
Dambacher, M., Dimigen, O., Braun, M., Wille, K., Jacobs, A.M., Kliegl, R., 2012. Stimulus onset asynchrony and the timeline of word recognition: event-related potentials during sentence reading. Neuropsychologia 50, 1852–1870.
Fischler, I., Bloom, P.A., Childers, D.G., Roucos, S.E., Perry, N.W., 1983. Brain potentials related to stages of sentence verification. Psychophysiology 20, 400–409.
Frazier, L., Carlson, K., Clifton, C., 2006. Prosodic phrasing is central to language comprehension. Trends Cogn. Sci. 10, 244–249.
Hagoort, P., Brown, C.M., 2000a. ERP effects of listening to speech compared to reading: the P600/SPS to syntactic violations in spoken sentences and rapid serial visual presentation. Neuropsychologia 38, 1531–1549.
Hagoort, P., Brown, C.M., 2000b. ERP effects of listening to speech: semantic ERP effects. Neuropsychologia 38, 1518–1530.
Holcomb, P.J., Neville, H.J., 1991. Natural speech processing: an analysis using event-related brain potentials. Psychobiology 19, 286–300.
Ito, A., Corley, M., Pickering, M.J., Martin, A.E., Nieuwland, M.S., 2016. Predicting form and meaning: evidence from brain potentials. J. Mem. Lang.
Kounios, J., Holcomb, P.J., 1992. Structure and process in semantic memory: evidence from event-related brain potentials and reaction times. J. Exp. Psychol. Gen. 121, 459.
Kutas, M., Federmeier, K.D., 2011. Thirty years and counting: finding meaning in the N400 component of the event related brain potential (ERP). Annu. Rev. Psychol., 621–647.
Kutas, M., Hillyard, S.A., 1980. Reading senseless sentences: brain potentials reflect semantic incongruity. Science 207, 203–205.
Kutas, M., Neville, H.J., Holcomb, P.J., 1987. A preliminary comparison of the N400 response to semantic anomalies during reading listening and signing. Electroencephalogr. Clin. Neurophysiol. Suppl. 39, 325–330.
Legge, G.E., Ahn, S.J., Klitz, T.S., Luebker, A., 1997. Psychophysics of reading—XVI. The visual span in normal and low vision. Vis. Res. 37, 1999–2010.
Marslen-Wilson, W.D., 1987. Functional parallelism in spoken word-recognition. Cognition 25, 71–102.
Nieuwland, M.S., 2016. Quantification, prediction and the online impact of sentence truth-value: evidence from event-related potentials. J. Exp. Psychol. Learn. Mem. Cogn.
Nieuwland, M.S., Kuperberg, G.R., 2008. When the truth is not too hard to handle an event-related potential study on the pragmatics of negation. Psychol. Sci. 19, 1213–1218.
Nieuwland, M.S., Martin, A.E., 2012. If the real world were irrelevant, so to speak: the role of propositional truth-value in counterfactual sentence comprehension. Cognition 122, 102–109.
Nieuwland, M.S., Van Berkum, J.J., 2006. When peanuts fall in love: N400 evidence for the power of discourse. J. Cognit. Neurosci. 18 (7), 1098–1111.
Öhman, S.E., 1966. Coarticulation in VCV utterances: spectrographic measurements. J. Acoust. Soc. Am. 39, 151–168.
Partee, B., 1991. Topics, focus, and quantification. In: Hastings, R., Jackson, B., Zvolenski, Z. (Eds.), Proceedings From Semantics and Linguistic Theory XI. Ithaca: Clc. pp. 159–187.
R Core Team, 2014. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL ⟨http://www.R-project.org/⟩.
Rayner, K., 1998. Eye movements in reading and information processing: 20 years of research. Psychol. Bull. 124, 372–422.
Tauroza, S., Allison, D., 1990. Speech rates in British English. Appl. Linguist. 11, 90–105.
Urbach, T.P., DeLong, K.A., Kutas, M., 2015. Quantifiers are incrementally interpreted in context, more than less. J. Mem. Lang. 83, 79–96.
Urbach, T.P., Kutas, M., 2010. Quantifiers more or less quantify on-line: ERP evidence for partial incremental interpretation. J. Mem. Lang. 63, 158–179.
Van Berkum, J.J., Zwitserlood, P., Hagoort, P., Brown, C.M., 2003. When and how do listeners relate a sentence to the wider discourse? Evidence from the N400 effect. Cognit. Brain Res. 17, 701–718.
Van Petten, C., Coulson, S., Rubin, S., Plante, E., Parks, M., 1999. Time course of word identification and semantic integration in spoken language. J. Exp. Psychol. Learn. Mem. Cogn. 25, 394–417.